

**Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (canceled) A method comprising:
  - (a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;
  - (b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and
  - (c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model.
2. (previously presented) The method of claim 4, wherein parts (a) – (c) are repeated at least once.
3. (previously presented) The method of claim 4, wherein the utterances are converted into a recognized phone string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.
4. (currently amended) A method comprising:
  - (a) calculating estimated weights for identified errors in recognition of utterances of a speaker based on a reference string, the utterances being received by a speaker input and converted to digital signals;
  - (b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and
  - (c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model;  
wherein calculating the estimated weights comprises computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.

5. (canceled) The method of claim 1, wherein calculating the estimated weights comprises computing an average likelihood difference per frame according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

6. (currently amended) A method comprising:

(a) calculating estimated weights for identified errors in recognition of utterances of a speaker based on a reference string, the utterances being received by a speaker input and converted to digital signals;

(b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and

(c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model;

wherein calculating the estimated weights comprises computing an average likelihood difference per frame according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string, and computing a weight for misrecognized words of a particular speaker “i” according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

7. (previously presented) The method of claim 4, wherein for a particular speaker, different misrecognized words have different weights.

8. (canceled) A method comprising:

- (a) recognizing utterances through converting the utterances into a recognized string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model.

9. (previously presented) The method of claim 12, wherein the utterances are converted into the recognized string through applying the speaker independent model.

10. (previously presented) The method of claim 12, wherein parts (b) – (e) are repeated until differences between the reference and recognized strings are less than a threshold.

11. (previously presented) The method of claim 12, wherein the utterances are converted into a recognized string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.

12. (currently amended) A method comprising:

- (a) recognizing utterances of a speaker through converting the utterances into a recognized string, the utterances being received by a speaker input and converted to digital signals;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and

(e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model;

wherein calculating the estimated weights comprises computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over all error words.

13. (canceled) The method of claim 8, wherein calculating the estimated weights comprises calculating an average likelihood difference per frame according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

14. (currently amended) A method comprising:

(a) recognizing utterances of a speaker through converting the utterances into a recognized string, the utterances being received by a speaker input and converted to digital signals;

(b) comparing the recognized string with a reference string to determine errors;

(c) calculating estimated weights for sections of the utterances;

(d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and

(e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model;

wherein calculating the estimated weights comprises calculating an average likelihood difference per frame according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string, and calculating a weight for misrecognized words of a particular speaker “i” is calculated according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |L_n| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

15. (previously presented) The method of claim 12, wherein for a particular speaker, different misrecognized words have different weights.

16. (canceled) An article of manufacture comprising:

a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

(a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;

(b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and

(c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model.

17. (previously presented) The article of manufacture of claim 19, wherein parts (a) – (c) are repeated at least once.

18. (previously presented) The article of manufacture of claim 19, wherein the utterances are converted into a recognized phone string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.

19. (previously presented) An article of manufacture comprising:

a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

(a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;

(b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and

(c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model;

wherein the estimated weights are computed through computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.

20. (canceled) The article of manufacture of claim 16, wherein an average likelihood difference per frame is used to calculate the estimated weights and is computed according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

21. (previously presented) An article of manufacture comprising:  
a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

(a) calculating estimated weights for identified errors in recognition of utterances based on a reference string;

(b) marking sections of the utterances as being misrecognized and associating the estimated weights with the sections of the utterances; and

(c) using the weighted sections of the utterances to convert a speaker independent model to a speaker dependent model;

wherein an average likelihood difference per frame is used to calculate the estimated weights and is computed according to equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string, and

a weight for misrecognized words of a particular speaker “i” is calculated according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |Ln| \quad (2), \text{ wherein } m$$

a number of misrecognized words.

22. (previously presented) The article of manufacture of claim 19, wherein for a particular speaker, different misrecognized words have different weights.

23. (canceled) An article of manufacture comprising:

a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

- (a) recognizing utterances through converting the utterances into a recognized phone string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model.

24. (previously presented) The article of manufacture of claim 27, wherein the utterances are converted into the recognized string through applying the speaker independent model.

25. (previously presented) The article of manufacture of claim 27, wherein parts (b) – (e) are repeated until differences between the reference and recognized strings are less than a threshold.

26. (previously presented) The article of manufacture of claim 27, wherein the utterances are converted into a recognized string a first time through applying the speaker independent model and thereafter through applying the most recently obtained speaker dependent model.

27. (previously presented) An article of manufacture comprising:  
a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

- (a) recognizing utterances through converting the utterances into a recognized phone string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model;

wherein the estimated weights are computed through computing an average likelihood difference per frame and then computing a weight value by averaging the average likelihood difference over error words.

28. (canceled) The article of manufacture of claim 23, wherein an average likelihood difference per frame is used to calculate the estimated weights and is calculated according to the equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_c^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string.

29. (previously presented) An article of manufacture comprising:

a computer-readable storage medium having executable instructions thereon which when executed cause a processor to perform operations comprising:

- (a) recognizing utterances through converting the utterances into a recognized phone string;
- (b) comparing the recognized string with a reference string to determine errors;
- (c) calculating estimated weights for sections of the utterances;
- (d) marking the errors in the utterances and providing corresponding estimated weights to form adaptation enrollment data; and
- (e) using the adaptation enrollment data to convert a speaker independent model to a speaker dependent model;

wherein an average likelihood difference per frame is used to calculate the estimated weights and is calculated according to the equation (1) as follows:

$$Ln = \frac{H_L^n}{H_e^n - H_b^n} - \frac{R_L^n}{R_e^n - R_b^n} \quad (1),$$

where  $H_L^n$  is a log likelihood of hypothesis word n,  $H_b^n$  is a beginning frame index (in time), and  $H_e^n$  is an end frame index, and  $R_L^n$ ,  $R_b^n$  and  $R_e^n$  are counter parts for the reference string, and

a weight for misrecognized words of a particular speaker "i" is calculated according to equation (2) as follows:

$$W_i = \frac{1}{m} * \sum_{n=1}^m |Ln| \quad (2), \text{ wherein } m$$

is a number of misrecognized words.

30. (previously presented) The article of manufacture of claim 27, wherein for a particular speaker, different misrecognized words have different weights.

31. (previously presented) The method of claim 4 wherein calculating the estimated weights further comprises:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

32 (currently amended) The method of claim 6 herein wherein calculating the estimated weights further comprises:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

33 (currently amended) The method of claim [[1]] 12 wherein calculating the estimated weights further comprises:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

34. (previously presented) The article of manufacture of claim 19 wherein the executable instructions causing the processor to perform calculating estimated weights comprises executable instructions thereon which when executed cause the processor to perform operations comprising:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

35. (currently amended) The article of manufacture of claim [[23]] 21 wherein the executable instructions causing the processor to perform calculating estimated weights comprises executable instructions thereon which when executed cause the processor to perform operations comprising:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

36. (previously presented) The article of manufacture of claim 27 wherein the executable instructions causing the processor to perform calculating estimated weights comprises executable instructions thereon which when executed cause the processor to perform operations comprising:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

37. (previously presented) The article of manufacture of claim 29 wherein the executable instructions causing the processor to perform calculating estimated weights comprises executable instructions thereon which when executed cause the processor to perform operations comprising:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and

aligning the 1-best hypothesis with the reference string to obtain the error words.

38 (new) The method of claim 14 wherein calculating the estimated weights further comprises:

running a force alignment program on the reference string to obtain statistics of references;

decoding the utterances to obtain statistics of 1-best hypothesis; and  
aligning the 1-best hypothesis with the reference string to obtain the error words.